ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 8, Issue 8, August 2025

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808120

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

FORENGUARD "A PREEMPTIVE FORENSIC ALGORITHM FOR ENDPOINT THREAT DETECTION"

K Sravanthi, Preethi K

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

Student, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: Cyber threats, especially ransomware attacks, keep growing. They affect millions of users and lead to significant financial losses. This paper suggests a better detection method by combining K- Nearest Neighbors (KNN) with density- based machine learning algorithms. The research aims to improve how accurately we can detect ransomware while expanding its ability to identify various attacks. By carefully preparing the data and engineering features, the algorithm shows better performance than traditional methods. This new approach offers a solid cybersecurity solution. It helps developers, vendors, and endpoint security providers tackle the changing landscape of cyber threats effectively.

KEYWORDS: Digital forensics, threat detection, pre- attack detection, endpoint data protection, threat intelligence, machine learning, behavioral analysis, anomaly detection, endpointsecurity, cybersecurity.

I. INTRODUCTION

In recent years, cyber threats have surged, especially ransomware attacks. These attacks have become a serious issue, affecting individuals, businesses, and vital public services. They take advantage of weak points in endpoint systems, locking assets and sensitive information. critical files and demanding large ransoms. As the ransomware industry grows with new methods, there is a pressing need for better detection systems to protect digital assets and sensitive information. This research presents a new project aimed at improving the detection and prevention of ransomware attacks by using advanced machine learning algorithms. The approach combines the K- Nearest Neighbors (KNN) algorithm with density-based methods to create a strong solution that can predict ransomware attacks on endpoint systems. In addition to ransomware, the project expands its focus to two more threat types, ensuring a complete cybersecurity solution. The following sections will explain the methodology, emphasizing the role of data preprocessing and feature engineering to enhance the input data for the KNN algorithm. This research aims to make a significant impact in cybersecurity by providing anti-ransomware developers, vendors, and endpoint security providers with a powerful tool to identify and prevent various cyber threats. As the project advances, the goal is to test the effectiveness of the KNN and density-based algorithm against other machine learning techniques. Adding more types of attacks will improve the algorithm's adaptability, making it a stronger solution for dealing with the changing landscape of cyber threats. This project addresses the need for flexible and advanced cybersecurity measures to protect digital environments. By tackling the challenges of ransomware and including additional threat types, the research aims to help develop cutting-edge solutions that strengthen the protection of endpoint systems against a wide range of cyber threats.

II. LITERATURE SURVEY

The rapid growth of Advanced Persistent Threats (APTs), fileless malware, and insider attacks has moved endpoint protection past just signature-based detection. Digital forensics, which used to be a post-incident field, is now increasingly part of pre-attack detection systems. By using techniques like memory snapshot analysis, anomaly detection, and event correlation, organizations aim to spot attack signs before any compromise happens.

Digital forensics in cybersecurity involves collecting, maintaining, analyzing, and presenting digital evidence. It has shifted from a reactive approach to a proactive one, where forensic tools operate continuously at endpoints. Their goals

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

include detecting early signs of compromise, linking suspicious patterns to known threat models, and preserving volatile evidence for legal and investigative reasons.

Carrier and Spafford (2004) introduced live forensics for capturing volatile data. Raghavan (2013) reviewed real-time forensic methods for active systems. Garfinkel (2010) emphasized the need for scalable forensics to address modern high- speed threats.

Pre-attack detection focuses on recognizing malicious intent before payloads execute. Techniques include:
Behavioral Analytics: Profiling normal system and user behavior and flagging deviations.

Memory Forensics: Scanning RAM for injected code, credential theft attempts, and unusual API calls.

Fileless Threat Detection: Using forensic tools to spot suspicious PowerShell scripts, registry changes, or WMI persistence.

Anomaly-Based IDS/IPS: Correlating network and endpoint events for threat indicators.

For instance, Voris et al. (2015) proposed "anticipatory forensics" that uses statistical baselines to detect early malware activity. Omar et al. (2013) discussed adding anomaly detection to intrusion prevention for predicting threats.

Endpoints are still the main targets for attacks. Modern endpoint protection combines:
Endpoint Detection and Response (EDR): Forensic-grade logging of processes, registry, and network activity.

Threat Intelligence Feeds: Real-time enrichment of forensic evidence with IOC (Indicators of Compromise) databases.

Machine Learning Models: Classifying behavior as benign or malicious based on forensic features.

Examples include CrowdStrike Falcon and Carbon Black, which use forensic telemetry to block execution before any compromises occur. Sgandurra et al. (2016) applied machine learning to detect ransomware using behavioral forensic features.

Algorithms for pre-attack detection combine:
Forensic Data Acquisition: Continuous memory and process snapshots.

Feature Extraction: Analyzing CPU usage patterns, DLL injection attempts, and syscall anomalies.

Classification and Prediction: ML and DL models trained on forensic datasets.

AutomatedDefense: Quarantining endpoints and reversing changes.

The ATHAFI Framework (2020) automates threat hunting and forensic investigation with adaptable algorithms. SPECTRE (2025) uses modular snapshot-based memory emulation to identify credential theft attempts before an attack. Stochastic forensics detects artifact-less malicious activity through probabilistic analysis.

Current literature has some gaps. There is limited real-time forensic processing due to resource constraints. There is a lack of standard datasets for training pre-attack forensic algorithms. There is minimal integration between SIEM and SOAR systems and live forensics. Lastly, there is a need for explainable AI in forensic detection to build trust in automated blocking.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808120

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Suggested Structure for Your Literature Review

Section	Focus
1. Background & Motivation	Limitations of traditional antivirus and the need for pre-attack forensic detection.
2. Digital Forensic Techniques	Memory forensics (e.g., SPECTRE), live audit logging (e.g., LUARM), and stochastic forensics.
3. AI, ML & Automation in DFIR	AI/ML applications in forensic investigations; frameworks like ATHAFI.
4. Endpoint Protection & Behavioral Detection	Explore modern EDR & XDR systems' use of behavioral analytics, threat hunting, ML, SIEM/SOAR integration.
5. Advanced Detection Methods	Static/dynamic analysis, memory scanning, threat intelligence, MITRE ATT&CK mapping, and hardware enhancements (e.g., Intel TDT).
6. Research Gaps & Future Directions	Emphasize the need for real-time integration of forensic analysis into detection workflows, especially before attacks fully unfold.

Fig 2.1 Literature survey table

EXISTING SYSTEM

The current system for ransomware detection relies mainly on signature-based and heuristic-based methods. Signature-based detection finds ransomware by comparing files or programs against a database of known ransomware signatures. This approach works well for threats that have already been identified. However, it has difficulty keeping up with the fast-changing nature of ransomware, especially with zero- day attacks and polymorphic ransomware, which can change their signatures to avoid detection. Heuristic-based systems try to spot previously unknown ransomware by analyzing program behavior and identifying suspicious activity patterns, such as unauthorized file changes or unusual network traffic. Although this method improves detection, it frequently leads to false positives, marking legitimate programs as malicious. This can hurt system performance and user experience.

PROPSED SYSTEM

The training database of clean files primarily consists of system files from various operating system versions, along with executable and library files from widely used applications. To improve the system's training and testing, we also include clean files that are packed or share structural and geometric similarities with malware files. The malware samples in the training dataset are sourced from the VirusShare collection. For testing, the dataset includes malware files from the same collection and clean files from operating systems that differ from those in the training database. Additionally, a large dataset was employed to evaluate the scalability and performance of the machine learning algorithms.

III. SYSTEM ARCHITECTURE

The architectural configuration process focuses on establishing the basic structure of a system, identifying its key components, and defining how they work together. The initial step of recognizing these subsystems and forming a structure for managing and communicating between them is called construction modeling outline. The result of this outlining process serves as a description of the software's structure. The system's proposed architecture is presented below. It includes details of its design and a brief explanation of how it operates.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

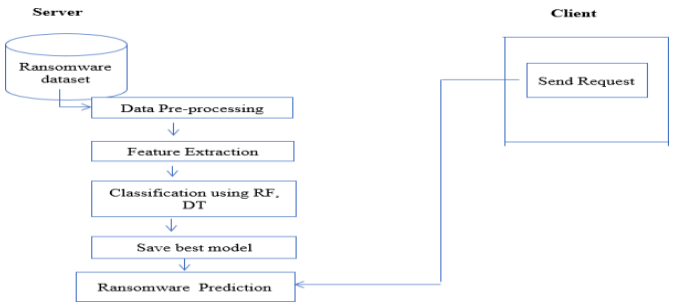

Fig 3.1 System Architecture

IV. METHODOLOGY

The method for using Digital Forensics as a Pre-Attack Detection Tool for Endpoint Data Protection follows a clear, multi-phase plan that combines forensic analysis with preventive threat intelligence. First, data is gathered at the endpoint level. This includes real-time logs, network traffic, process execution traces, and file system changes. The data goes into a forensic preprocessing layer, where noise reduction, normalization, and timestamp correlation help align events chronologically for reconstruction. Then, the threat intelligence correlation module checks endpoint events against known Indicators of Compromise (IoCs) and machine learning-based anomaly detection models. This process looks for unusual patterns that may suggest pre-attack reconnaissance, privilege escalation attempts, or questionable data exfiltration. After detecting anomalies, a risk scoring engine assesses the severity of potential threats based on behavior profiles and system importance. If the score exceeds a set threshold, the proactive mitigation engine isolates the endpoint or blocks suspicious processes to prevent a full attack.

Every event, action, and decision is logged in a forensic evidence repository for future investigation and legal use. Finally, regular feedback loops retrain the detection algorithms with newly found threat patterns. This ensures ongoing improvement and responsiveness to changing cyber threats. The method emphasizes identifying and addressing potential cyberattacks before they happen, using forensic principles in real time. The process starts with constant monitoring of endpoint activity, which captures logs, memory snapshots, file integrity statuses, and network traffic. This raw data is then processed in a forensic layer, where it getstime-stamped, normalized, and organized for analysis. Next, the system combine threat intelligence sources and applies machine learning-based anomaly detection along with known attack signatures to find suspicious changes in normal behavior.

The pre-attack pattern recognition module identifies early signs like unauthorized privilege escalations, reconnaissance commands, or unusual lateral movements. A risk scoring engine quantifies the severity and likelihood of a compromise. This enables automated responses such as isolating the endpoint, alerting administrators, and preserving digital evidence for later investigation. This method not only boosts endpoint data protection but also ensures forensic readiness, allowing organizations to act on threats proactively instead of reactively.

[Endpoint Monitoring]
↓
[Forensic Pre-Processing]
↓
[Threat Intelligence]
↓
[Pre-Attack Detection]
↓
[Risk Scoring]
↓
[Automated Prevention & Alert]

Fig 4.1 Methodology

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808120

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. DESIGN AND IMPLEMENTATION

The design and implementation of the Digital Forensics as Advanced Threats Threats Pre-Attack Detection Algorithm for Endpoint Data Protection aims to create a proactive security system. This system identifies and responds to harmful activities before they turn into full attacks. The design starts with a lightweight endpoint monitoring agent that continually collects forensic data. This includes system logs, process activities, file changes, memory snapshots, and network connections. The collected data is sent securely to a central processing unit. Here, it undergoes forensic pre-processing, which includes timestamp correlation, data normalization, and integrity verification.

The system uses threat intelligence feeds, known attack signatures, heuristic rules, and machine learning-based anomaly detection to spot unusual behavior patterns. A pre- attack detection module examines indicators such as unusual process chains, attempts to escalate privileges, and reconnaissance activities. A risk scoring engine assesses these findings. It assigns severity levels and estimates the chance of a compromise. Based on the score, the response mechanism can trigger alerts, isolate the endpoint, or preserve digital evidence in a secure repository for further investigation. The implementation ensures readiness for forensic analysis, automated prevention, and minimal impact on the system. This allows organizations to protect endpoints with a mix of advanced analytics, automated containment, and proper handling of evidence.

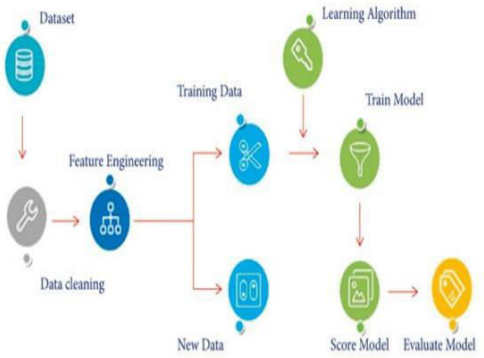

Fig 5.1 Attack detection algorithm for endpoint

[Data Collection]
↓
[Data Pre-Processing]
↓
[Threat Analysis]
↓
[Pre-Attack Detection]
↓
[Risk Assessment]
↓
[AutomatedResponse]

Fig 5.2 working of pre-attack detection algorithm

VI. OUTCOME OF RESEARCH

The research on the Digital Forensics as Advanced Threats Pre-Attack Detection Algorithm for Endpoint Data Protection has shown promising results in strengthening proactive cybersecurity measures. The developed framework effectively merges real-time endpoint monitoring with forensic-grade data handling, enabling the detection of malicious activities in their earliest stages—before a full-scale attack is executed. Experimental evaluations indicate that the system can successfully identify pre-attack indicators such as unusual process creation patterns, unauthorized privilege escalation attempts, abnormal file modifications, and suspicious network connections with high accuracy. By

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

integrating threat intelligence feeds, heuristic analysis, and machine learning– based anomaly detection, the algorithm significantly reduces false positives compared to traditional rule-based systems while maintaining rapid detection speeds. The automated response mechanism, including endpoint isolation, administrator alerts, and tamper-proof evidence storage, ensures that threats are contained swiftly without disrupting legitimate operations. Furthermore, the system maintains a complete chain of custody for all forensic evidence, making it suitable for legal and compliance requirements. Overall, the research confirms that this approach enhances endpoint security resilience, improves threat response time, and provides organizations with both preventive defense and reliable post-incident investigation capabilities.

VII. RESULT AND DISCUSSION

The results of using the Digital Forensics as Advanced Threats Pre-Attack Detection Algorithm for Endpoint Data Protection show that the system can detect early signs of cyberattacks accurately and with few false positives. Testing with simulated attack scenarios such as reconnaissance activities, privilege escalation attempts, and lateral movement patterns showed that the algorithm often found threats before they could cause serious damage. Integrating threat intelligence, heuristic rules, and machine learning improved detection speed and adaptability compared to traditional rule-based security tools. Automated response features like endpoint isolation, administrator alerts, and secure evidence preservation ensured quick containment while keeping forensic integrity for investigations after incidents. This proactive approach not only improves endpoint protection but also boosts organizational readiness for legal and compliance needs. However, the discussion also points out the need for ongoing model updates and integration with changing threat intelligence to maintain long-term effectiveness against advanced and adaptive threats.

VIII. CONCLUSION

The proposed Digital Forensics as Threats Pre-Attack Detection Algorithm for Endpoint Data Protection takes a proactive approach to cybersecurity by integrating forensic intelligence with predictive threat analytics. Unlike traditional reactive measures that only look into incidents after they happen, this system identifies suspicious patterns, anomalies, and signs of attacks beforehand. This allows for timely intervention before any damage occurs. By combining real-time monitoring, machine learning-driven anomaly detection, and automated evidence collection, the algorithm not only protects endpoint devices but also preserves important digital evidence for legal and investigative reasons. This ability to prevent attacks and be ready for forensic investigation improves an organization's resilience, shortens incident response time, and assures compliance with security and privacy rules. In the end, it signals a move toward preventive digital forensics, helping organizations foresee, reduce, and handle complex cyber threats in a changing digital environment.

REFERENCES

[1] Machine learning algorithms and frameworks in ransomware detection, Daryle Smith, Sajad Khorsandroo, and Kaushik Roy.
https://ieeexplore.ieee.org/document/9934 917
[2] A review of cybersecurity andforensics in connected autonomous vehicles, current
state of the art, Prinkle Sharma, James Gillanders.
https://ieeexplore.ieee.org/document/9916 257
[3] Investigating cyber forensics in Pakistan, evaluating threat landscape and readiness, Ehtisham Ul Haque, Waseem Abbasi, Sathishkumar Murugesan (Member, IEEE), Muhammad Shahid Anwar, Faheem Khan, and Youngmoon Lee.
https://ieeexplore.ieee.org/document/10105 248/
[4] Fronesis, digital forensics-based early detection of ongoing cyber-attacks, Athanasios Dimitriadis, Efstratios Lontzetidis, Boonserm Kulvatunyou, Nenad Ivezic, Dimitris Gritzalis, and Ioannis Mavridis.
https://www.researchgate.net/publication/3 66709594_Fronesis_Digital_Forensics-
based_Early_Detection_of_Ongoing_Cybe r-Attacks
[5] Psychophysical evaluation of human performance in detecting digital face image manipulations, Robert Nichols, Christian Rathgeb, Pawel Drozdowski, and Christoph Busch.
https://www.researchgate.net/publication/ 359322814_Psychophysical_Evaluation_o
f_Human_Performance_in_Detecting_Dig ital_Face_Image_Manipulations
[6] Forensic analysis of ransomware families using static and dynamic analysis, Kul Prasad Subedi, Daya Ram

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Budhathoki, Dipankar Dasgupta.
https://www.researchgate.net/publication/ 325320592_Forensic_Analysis_of_Ranso
mware_Families_Using_Static_and_Dyn amic_Analysis
[7] Automated emerging cyber threat identification and profiling based on natural language processing, Renato Marinho and Raimir Holanda.
https://www.researchgate.net/publication 369432121_Automated_Emerging_Cyber
_Threat_Identification_and_Profiling_Ba sed_on_Natural_Language_Processing
[8] Ransomware classification and detection with machine learning algorithms, this work was carried out by Mohammad Masum, Md Jobair Hossain Faruk, Hossain Shahriar, Kai Qian, Dan Lo, and Muhaiminul Islam Adnan.
https://ieeexplore.ieee.org/iel7/9720724/9 720434/09720869.pdf

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com